

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Eurogene Limited
- (B) STREET: Marquis House, 67/68 Jermyn Street
- (C) CITY: London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): SW1Y 6NY

(ii) TITLE OF INVENTION: BIOTIN-BINDING RECEPTOR MOLECULES

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5177 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1071..2270

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TTTGAAAGAC CCCACCCGTA GGTGGCAAGC TAGCTTAAGT AACGCCACTT TGCAAGGCAT	60
GGAAAAATAC ATAACTGAGA ATAGAAAAGT TCAGATCAAG GTCAGGAACA AAGAAACAGC	120
TGAATACCAA ACAGGATATC TGTGGTAAGC GGTTCCTGCC CCGGCTCAGG GCCAAGAACA	180
GATGAGACAG CTGAGTGATG GGCCAAACAG GATATCTGTG GTAAGCAGTT CCTGCCCCGG	240
CTCGGGGGCCA AGAACAGATG GTCCCCAGAT GCGGTCCAGC CCTCAGCAGT TTCTAGTGAA	300
TCATCAGATG TTTCCAGGGT GCCCCAAGGA CCTGAAAATG ACCCTGTACC TTATTTGAAC	360
TAACCAATCA GTTCGCTTCT CGCTTCTGTT CGCGCGCTTC CGCTCTCCGA GCTCAATAAA	420
AGAGCCCACA ACCCCTCACT CGGCGCGCCA GTCTTCCGAT AGACTGCGTC GCCCGGGTAC	480
CCGTATTCCC AATAAAGCCT CTTGCTGTTT GCATCCGAAT CGTGGTCTCG CTGTTCTTGT	540
GGAGGGTCTC CTCTGAGTGA TTGACTACCC ACGACGGGGG TCTTTCATTT GGGGGCTCGT	600
CCGGGATTTG GAGACCCCTG CCCAGGGACC ACCGACCCAC CACCGGGAGG TAAGCTGGCC	660
AGCAACTTAT CTGTGTCTGT CCGATTGTCT AGTGTCTATG TTTGATGTTA TGCGCCTGCG	720
TCTGTACTAG TTAGCTAACT AGCTCTGTAT CTGGCGGACC CGTGGTGGA CTGACGAGTT	780
CTGAACACCC GGCCGCAACC CTGGGAGACG TCCCAGGGAC TTTGGGGGCC GTTTTGTGG	840
CCCACCTGA GGAAGGGAGT CGATGTGGAA TCCGACCCCG TCAGGATATG TGGTCTCTGGT	900
AGGAGACGAG AACCTAAAAC AGTTCCCGCC TCCGTCTGAA TTTTGTCTTT CGGTTTGGA	960
CCGAAGCCGC GCGTCTTGTC TGCTGCAGCC AAGCTTGGGC TGCAGGTCGA CTCTAGAGGA	1020

TCAATTCGGC ACGAGTAAAT CGGTGCTGCC GTCTTTAGGA CATATGAAGT ATG GCA	1076
Met Ala	
1	
CAG TGG GAT GAC TTT CCT GAT CAG CAA GAG GAC ACT GAC AGC TGT ACA	1124
Gln Trp Asp Asp Phe Pro Asp Gln Gln Glu Asp Thr Asp Ser Cys Thr	
5 10 15	
GAG TCT GTG AAG TTC GAT GCT CGC TCA GTG ACA GCT TTG CTT CCT CCC	1172
Glu Ser Val Lys Phe Asp Ala Arg Ser Val Thr Ala Leu Leu Pro Pro	
20 25 30	
CAT CCT AAA AAT GGC CCA ACT CTT CAA GAG AGG ATG AAG TCT TAT AAA	1220
His Pro Lys Asn Gly Pro Thr Leu Gln Glu Arg Met Lys Ser Tyr Lys	
35 40 45 50	
ACT GCA CTG ATC ACC CTT TAT CTC ATT GTG TTT GTA GTT CTC GTG CCC	1268
Thr Ala Leu Ile Thr Leu Tyr Leu Ile Val Phe Val Val Leu Val Pro	
55 60 65	
ATC ATT GGC ATA GTG GCA GCT CAG CTC CTG AAA TGG GAA ACG AAG AAT	1316
Ile Ile Gly Ile Val Ala Ala Gln Leu Leu Lys Trp Glu Thr Lys Asn	
70 75 80	
TGC ACG GTT GGC TCA GTT AAT GCA GAT ATA TCT CCA AGT CCG GAA GGC	1364
Cys Thr Val Gly Ser Val Asn Ala Asp Ile Ser Pro Ser Pro Glu Gly	
85 90 95	
AAA GGA AAT GGC AGT GAA GAT GAA ATG AGA TTT CGA GAA GCT GTG ATG	1412
Lys Gly Asn Gly Ser Glu Asp Glu Met Arg Phe Arg Glu Ala Val Met	
100 105 110	
GAA CGC ATG AGC AAC ATG GAA AGC AGA ATC CAG TAT CTT TCA GAT AAT	1460
Glu Arg Met Ser Asn Met Glu Ser Arg Ile Gln Tyr Leu Ser Asp Asn	
115 120 125 130	

GAA GCC AAT CTC CTA GAT GCT AAG AAT TTC CAA AAT TTC AGC ATA ACA	1508
Glu Ala Asn Leu Leu Asp Ala Lys Asn Phe Gln Asn Phe Ser Ile Thr	
135 140 145	
ACT GAT CAA AGA TTT AAT GAT GTT CTT TTC CAG CTA AAT TCC TTA CTT	1556
Thr Asp Gln Arg Phe Asn Asp Val Leu Phe Gln Leu Asn Ser Leu Leu	
150 155 160	
TCC TCC ATC CAG GAA CAT GAG AAT ATC ATA GGG GAT ATC TCC AAG TCA	1604
Ser Ser Ile Gln Glu His Glu Asn Ile Ile Gly Asp Ile Ser Lys Ser	
165 170 175	
TTA GTA GGT CTG AAC ACC ACA GTA CTT GAT TTG CAG TTC AGT ATT GAA	1652
Leu Val Gly Leu Asn Thr Thr Val Leu Asp Leu Gln Phe Ser Ile Glu	
180 185 190	
ACA CTG AAT GGC AGA GTC CAA GAG AAT GCA TTT AAA CAA CAA GAG GAG	1700
Thr Leu Asn Gly Arg Val Gln Glu Asn Ala Phe Lys Gln Gln Glu Glu	
195 200 205 210	
ATG CGT AAA TTA GAG GAG CGT ATA TAC AAT GCA TCA GCA GAA ATT AAG	1748
Met Arg Lys Leu Glu Glu Arg Ile Tyr Asn Ala Ser Ala Glu Ile Lys	
215 220 225	
TCT CTA GAT GAA AAA CAA GTA TAT TTG GAA CAG GAA ATA AAA GGG GAA	1796
Ser Leu Asp Glu Lys Gln Val Tyr Leu Glu Gln Glu Ile Lys Gly Glu	
230 235 240	
ATG AAA CTG TTG AAT AAT ATC ACT AAT GAT CTG AGG CTG AAG GAT TGG	1844
Met Lys Leu Leu Asn Asn Ile Thr Asn Asp Leu Arg Leu Lys Asp Trp	
245 250 255	
GAA CAT TCT CAG ACA TTG AAA AAT ATC ACT TTA CTC CAA GGT GCC AGA	1892
Glu His Ser Gln Thr Leu Lys Asn Ile Thr Leu Leu Gln Gly Ala Arg	
260 265 270	

AAG TGC TCG CTG ACT GGG AAA TGG ACC AAC GAT CTG GGC TCC AAC ATG	1940
Lys Cys Ser Leu Thr Gly Lys Trp Thr Asn Asp Leu Gly Ser Asn Met	
275 280 285 290	
ACC ATC GGG GCT GTG AAC AGC AGA GGT GAA TTC ACA GGC ACC TAC ATC	1988
Thr Ile Gly Ala Val Asn Ser Arg Gly Glu Phe Thr Gly Thr Tyr Ile	
295 300 305	
ACA GCC GTA ACA GCC ACA TCA AAT GAG ATC AAA GAG TCA CCA CTG CAT	2036
Thr Ala Val Thr Ala Thr Ser Asn Glu Ile Lys Glu Ser Pro Leu His	
310 315 320	
GGG ACA CAA AAC ACC ATC AAC AAG AGG ACC CAG CCC ACC TTT GGC TTC	2084
Gly Thr Gln Asn Thr Ile Asn Lys Arg Thr Gln Pro Thr Phe Gly Phe	
325 330 335	
ACC GTC AAT TGG AAG TTT TCA GAG TCC ACC ACT GTC TTC ACG GGC CAG	2132
Thr Val Asn Trp Lys Phe Ser Glu Ser Thr Thr Val Phe Thr Gly Gln	
340 345 350	
TGC TTC ATA GAC AGG AAT GGG AAG GAG GTC CTG AAG ACC ATG TGG CTG	2180
Cys Phe Ile Asp Arg Asn Gly Lys Glu Val Leu Lys Thr Met Trp Leu	
355 360 365 370	
CTG CGG TCA AGT GTT AAT GAC ATT GGT GAT GAC TGG AAA GCT ACC AGG	2228
Leu Arg Ser Ser Val Asn Asp Ile Gly Asp Asp Trp Lys Ala Thr Arg	
375 380 385	
GTC GGC ATC AAC ATC TTC ACT CGC CTG CGC ACA CAG AAG GAG	2270
Val Gly Ile Asn Ile Phe Thr Arg Leu Arg Thr Gln Lys Glu	
390 395 400	
TGAGTGAGTG ACCAAGGTCC TCCTGGACTC CAGGTGAAAA AGGAGATAGA GGCCCTCCTG	2330
GACAAAATGG TATACCAGGC TTTCCAGGTC TAATAGGTAC TCCAGGTCTT AAAGGTGATC	2390
GGGGGGATCT CTGGTTTACC TGGAGTTCGA GGATTCCCAG GACCAATGGG GAAGACCGGG	2450

AAGCCAGGAC TTAATGGACA AAAAGGCCAG AAGGGAGAAA AAGGGAGTGG AAGCATGCAA	2510
AGACAATCTA ATACAGTCCG ACTGGTGGGT GGCAGCGGCC CTCACGAAGG CAGAGTGGAG	2570
ATTTTTCACG AAGGCCAGTG GGTACGGTG TGTGACGACC GCTGGGAACT GCGTGGAGGA	2630
CTGGTCGTCT GCAGGAGCTT GGGATACAAA GGTGTTCAAA GTGTGCATAA GCGAGCTTAT	2690
TTTGGAAGG GTACGGGTCC AATATGGCTG AATGAAGTAT TTTGTTTCGG GAAAGAGTCA	2750
TCCATTGAAG AGTGCAGAAT TAGACAGTGG GGTGTGAGAG CCTGTTCGCA CGACGAAGAT	2810
GCTGGGGGTC ACTTTGCACC TACATAATGC ATCATATTTT CATTCACATT TTTTAACTG	2870
TTATAAGTG ATTTTTTTCC TTTGCTTCAC TAAAATCAGC TTAATTAATA TTTAAGAAAC	2930
TAAGAATTTT ATCCACAGAA AAGGAATATT TAAAAATCAC TGGATAAACA TATAAAATAG	2990
CTTCATATTT GCTTCAAATA CCAGAACCAT TTCAACTTCT CTAGGTTTTT AAGTGGCTCG	3050
TGCCGAATTG ATCCCCTCAG GATATAGTAG TTTCGCTTTT GCATAGGGAG GGGGAAATGT	3110
AGTCTTATGC AATACTCTTG TAGTCTTGCA ACATGGTAAC GATGAGTTAG CAACATGCCT	3170
TACAAGGAGA GAAAAGCAC CGTGCATGCC GATTGGTGGG AGTAAGGTGG TACGATCGTG	3230
CCTTATTAGG AAGGCAACAG ACGGGTCTGA CATGGATTGG ACGAACCCT GAATTCCGCA	3290
TTGCAGAGAT ATTGTATTTA AGTGCCTAGC TCGATACAGC AAACGCCATT TGACCATTCA	3350
CCACATTGGT GTGCACCTCC AAGCTTCACG CTGCCGCAAG CACTCAGGGC GCAAGGGCTG	3410
CTAAAGGAAG CGGAACACGT AGAAAGCCAG TCCGCAGAAA CGGTGCTGAC CCCGGATGAA	3470
TGTCAGCTAC TGGGCTATCT GGACAAGGGA AAACGCAAGC GCAAAGAGAA AGCAGGTAGC	3530
TTGCAGTGGG CTTACATGGC GATAGCTAGA CTGGGCGGTT TTATGGACAG CAAGCGAACC	3590

GGAATTGCCA GCTGGGGCGC CCTCTGGTAA GGTGGGAAG CCCTGCAAAG TAAACTGGAT	3650
GGCTTTCTTG CCGCCAAGGA TCTGATGGCG CAGGGGATCA AGATCTGATC AAGAGACAGG	3710
ATGAGGATCG TTTCGCATGA TTGAACAAGA TGGATTGCAC GCAGGTTCTC CGGCCGCTTG	3770
GGTGGAGAGG CTATTCGGCT ATGACTGGGC ACAACAGACA ATCGGCTGCT CTGATGCCGC	3830
CGTGTTCGGG CTGTCAGCGC AGGGGCGCCC GGTTCTTTTT GTCAAGACCG ACCTGTCCGG	3890
TGCCCTGAAT GAACTGCAGG ACGAGGCAGC GCGGCTATCG TGGCTGGCCA CGACGGGCGT	3950
TCCTTGCGCA GCTGTGCTCG ACGTTGTCAC TGAAGCGGGA AGGGACTGGC TGCTATTGGG	4010
CGAAGTGCCG GGGCAGGATC TCCTGTCATC TCACCTTGCT CCTGCCGAGA AAGTATCCAT	4070
CATGGCTGAT GCAATGCGGC GGCTGCATAC GCTTGATCCG GCTACCTGCC CATTCGACCA	4130
CCAAGCGAAA CATCGCATCG AGCGAGCAGC TACTCGGATG GAAGCCGGTC TTGTCGATCA	4190
GGATGATCTG GACGAAGAGC ATCAGGGGCT CGCGCCAGCC GAACTGTTCG CCAGGCTCAA	4250
GGCGCGCATG CCCGACGGCG AGGATCTCGT CGTGACCCAT GGCGATGCCT GCTTGCCGAA	4310
TATCATGGTG GAAAATGGCC GCTTTTCTGG ATTCATCGAC TGTGGCCGGC TGGGTGTGGC	4370
GGACCGCTAT CAGGACATAG CGTTGGCTAC CCGTGATATT GCTGAAGAGC TTGGCGGCGA	4430
ATGGGCTGAC CGCTTCCTCG TGCTTTACGG TATCGCCGCT CCCGATTCGC AGCGCATCGC	4490
CTTCTATCGC CTTCTTGACG AGTTCTTCTG AGCGGGACTC TGGGGTTCGA TAAAATAAAA	4550
GATTTTATTT AGTCTCCAGA AAAAGGGGGG AATGAAAGAC CCCACCTGTA GGTTCGGCAA	4610
GCTAGCTTAA GTAACGCCAT TTTGCAAGGC ATGGAAAAAT ACATAACTGA GAATAGAGAA	4670
GTTTCAGATCA AGGTCAGGAA CAGATGGAAC AGCTGAATAT GGGCCAAACA GGATATCTGT	4730

GGTAAGCAGT TCCTGCCCCG GCTCAGGGCC AAGAACAGAT GGAACAGCTG AATATGGGCC 4790
 AAACAGGATA TCTGTGGTAA GCAGTTCCTG CCCC GGCTCA GGGCCAAGAA CAGATGGTCC 4850
 CCAGATGCGG TCCAGCCCTC AGCAGTTTCT AGAGAACCAT CAGATGTTTC CAGGGTGCCC 4910
 CAAGGACCTG AAATGACCCT GTGCCTTATT TGAAC TAACC AATCAGTTCG CTTCTCGCTT 4970
 CTGTTGCGGC GCTTCTGCTC CCCGAGCTCA ATAAAAGAGC CCACAACCCC TCACTCGGGG 5030
 CGCCAGTCCT CCGATTGACT GAGTCGCCCC GGTACCCGTG TATCCAATAA ACCCTCTTGC 5090
 AGTTGCATCC GACTTGTGGT CTCGCTGTTC CTTGGGAGGG TCTCCTCTGA GTGATTGACT 5150
 ACCCGTCAGC GGGGGTCTTT CATTGG 5177

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Gln Trp Asp Asp Phe Pro Asp Gln Gln Glu Asp Thr Asp Ser
 1 5 10 15
 Cys Thr Glu Ser Val Lys Phe Asp Ala Arg Ser Val Thr Ala Leu Leu
 20 25 30
 Pro Pro His Pro Lys Asn Gly Pro Thr Leu Gln Glu Arg Met Lys Ser
 35 40 45
 Tyr Lys Thr Ala Leu Ile Thr Leu Tyr Leu Ile Val Phe Val Val Leu
 50 55 60

Val Pro Ile Ile Gly Ile Val Ala Ala Gln Leu Leu Lys Trp Glu Thr
65 70 75 80

Lys Asn Cys Thr Val Gly Ser Val Asn Ala Asp Ile Ser Pro Ser Pro
85 90 95

Glu Gly Lys Gly Asn Gly Ser Glu Asp Glu Met Arg Phe Arg Glu Ala
100 105 110

Val Met Glu Arg Met Ser Asn Met Glu Ser Arg Ile Gln Tyr Leu Ser
115 120 125

Asp Asn Glu Ala Asn Leu Leu Asp Ala Lys Asn Phe Gln Asn Phe Ser
130 135 140

Ile Thr Thr Asp Gln Arg Phe Asn Asp Val Leu Phe Gln Leu Asn Ser
145 150 155 160

Leu Leu Ser Ser Ile Gln Glu His Glu Asn Ile Ile Gly Asp Ile Ser
165 170 175

Lys Ser Leu Val Gly Leu Asn Thr Thr Val Leu Asp Leu Gln Phe Ser
180 185 190

Ile Glu Thr Leu Asn Gly Arg Val Gln Glu Asn Ala Phe Lys Gln Gln
195 200 205

Glu Glu Met Arg Lys Leu Glu Glu Arg Ile Tyr Asn Ala Ser Ala Glu
210 215 220

Ile Lys Ser Leu Asp Glu Lys Gln Val Tyr Leu Glu Gln Glu Ile Lys
225 230 235 240

Gly Glu Met Lys Leu Leu Asn Asn Ile Thr Asn Asp Leu Arg Leu Lys
245 250 255

Asp Trp Glu His Ser Gln Thr Leu Lys Asn Ile Thr Leu Leu Gln Gly
260 265 270

Ala Arg Lys Cys Ser Leu Thr Gly Lys Trp Thr Asn Asp Leu Gly Ser
 275 280 285

Asn Met Thr Ile Gly Ala Val Asn Ser Arg Gly Glu Phe Thr Gly Thr
 290 295 300

Tyr Ile Thr Ala Val Thr Ala Thr Ser Asn Glu Ile Lys Glu Ser Pro
 305 310 315 320

Leu His Gly Thr Gln Asn Thr Ile Asn Lys Arg Thr Gln Pro Thr Phe
 325 330 335

Gly Phe Thr Val Asn Trp Lys Phe Ser Glu Ser Thr Thr Val Phe Thr
 340 345 350

Gly Gln Cys Phe Ile Asp Arg Asn Gly Lys Glu Val Leu Lys Thr Met
 355 360 365

Trp Leu Leu Arg Ser Ser Val Asn Asp Ile Gly Asp Asp Trp Lys Ala
 370 375 380

Thr Arg Val Gly Ile Asn Ile Phe Thr Arg Leu Arg Thr Gln Lys Glu
 385 390 395 400